

PRIMERGY BX900 / BX400 InfiniBand Modules (40Gb)

Connection Blade and Mezzanine Card V7.0
(Linux®, Windows®)

Comments... Suggestions... Corrections...

The User Documentation Department would like to know your opinion of this manual. Your feedback helps us optimize our documentation to suit your individual needs.

Feel free to send us your comments by e-mail to manuals@ts.fujitsu.com.

Certified documentation according to DIN EN ISO 9001:2008

To ensure a consistently high quality standard and user-friendliness, this documentation was created to meet the regulations of a quality management system which complies with the requirements of the standard DIN EN ISO 9001:2008.

cognitas. Gesellschaft für Technik-Dokumentation mbH
www.cognitas.de

Copyright and Trademarks

Copyright © Fujitsu Technology Solutions GmbH 2012.

All rights reserved.

Delivery subject to availability; right of technical modifications reserved.

All hardware and software names used are trademarks of their respective manufacturers.

Contents

1	Introduction	5
1.1	Notational Conventions	5
1.2	Target Group	6
1.3	Documentation Overview	6
1.4	Overview	7
1.5	PRIMERGY BX900	8
1.6	Inserting Connection Blade into PRIMERGY BX900	9
1.7	Connecting Mezzanine Cards on Server Blade	10
1.8	Description of InfiniBand Connection Blade	11
1.8.1	Cables	11
1.8.2	LEDs	12
1.9	Technical Data	12
1.9.1	IB Connection Blade	12
1.9.2	IB Mezzanine Card	13
2	Linux Installation	15
2.1	Overview	15
2.2	Software Requirements	16
2.3	Native InfiniBand Installation in RHEL	17
2.3.1	Setup InfiniBand during Linux Installation	17
2.3.2	Loading InfiniBand Drivers at System Startup	18
2.3.3	Installation Results	19
2.3.4	Post-installation Notes	19
2.4	OFED Installation by means of SVIM	20
2.4.1	In the event of RHEL5.4	20
2.4.2	In the event of RHEL6.0	24
2.4.3	Post-installation Notes	25
2.5	Updating Firmware after Installation	27
2.6	Subnet Manager	27
2.6.1	OpenSM	27

Contents

2.6.2	Confirmation of OpenSM running	28
2.6.3	Confirmation of the master SM	28
2.7	Network Configuration Information	28
2.7.1	Network configuration diagram	29
2.7.2	Configuration sheet	29
2.7.3	The result of the command "ibnetdiscover"	32
2.8	Identifying the GUID	35
3	Windows Installation	37
3.1	Overview	37
3.1.1	Software Requirements	37
3.2	Installation of WinOF	38
4	IB Connection Blade Management	51
4.1	Connection	51
4.1.1	telnet connection	52
4.1.2	Save and restore the configuration	52
4.2	Configuration	54
4.2.1	Configuration of LAN interface	54
4.2.2	Ethernet Connection via Management Blade	57
4.3	Basic Operation	58
4.3.1	How to check System Information	58
4.3.2	SNMP Trap Settings	59
4.4	Users and Privileges	60
5	Troubleshooting	61
5.1	Check the Log File of OpenSM (Linux only)	61
5.2	IB Mezzanine Card	64
5.3	IB Connection Blade	76
Index		79

1 Introduction

This manual describes the configuration, installation and basic use of the InfiniBand connection blade and Mezzanine card software. The package includes tools for general and advanced management functionality.

It also describes the installation of OFED (OpenFabrics Enterprise Distribution) for Linux, the identification of the Mezzanine card and Win OpenFabrics installation for Windows.

The hardware installation of the InfiniBand connection blade is described in the BX900 S1 or BX400 S1 System Unit Operation manual. The hardware installation of the Mezzanine card is described in the Service Supplements or Upgrade and Maintenance Manual of the BX92x S3 server blades.

1.1 Notational Conventions

The following notational conventions are used in this manual:

	Caution This symbol points out hazards that can lead to personal injury, loss of data, or damage to equipment.
	This symbol highlights important information and tips.
►	This symbol refers to a step that you must carry out in order to continue with the procedure.
<i>italic</i>	Commands, menu items, names of buttons, options, file names, and path names are written in <i>italic</i> letters in the text.
<variable>	Marked variables that must be replaced by current values.
fixed font	System output is written using a fixed font.
semi-bold fixed font	Commands to be entered through the keyboard are written in a semi-bold fixed font.
Key symbols	Keys are presented according to their representation on the keyboard. If capital letters are to be entered explicitly, then the Shift key is shown, e.g. [SHIFT] - [A] for A. If two keys need to be pressed at the same time, then this is indicated by placing a hyphen between the two key symbols.

Fujitsu Support

Please contact your Fujitsu Technical Support if you require assistance:

<http://support.ts.fujitsu.com/com/support/index.html>

If you purchased the products from Fujitsu Japan, please consult your system engineer.

1.2 Target Group

This manual is intended for users, developers, and system administrators responsible for setting up and maintaining switch system platforms using InfiniBand fabrics and networks. The switch system platforms must have a management server or client, in order for the package to work.

The manual assumes familiarity with the InfiniBand® Architecture Specification as well as Ethernet Architecture Specification.

1.3 Documentation Overview

 PRIMERGY manuals are available in PDF format on the ServerView Suite DVD 2. The ServerView Suite DVD 2 is part of the ServerView Suite supplied with every server.

If you no longer have the ServerView Suite DVDs, you can obtain the relevant current version using the order number U15000-C289 (the order number for Fujitsu Japan: please refer to the configurator of the server <http://jp.fujitsu.com/platform/server/primergy/system>.

The PDF files of the manuals can also be downloaded free of charge from the internet. The overview page showing the online documentation available on the internet can be found using the URL (for EMEA market): <http://manuals.ts.fujitsu.com>. The PRIMERGY server documentation can be accessed using the *Industry standard servers* navigation option.

If you purchased the products from Fujitsu Japan, please use the URL: <http://jp.fujitsu.com/platform/server/primergy/manual>.

1.4 Overview

This manual explains the use of the software management. Managed devices must have a CPU and a management board. This board has an operating system and a toolset to operate, manage, and upgrade the device.

Further information is provided on the PRIMERGY ServerBooks DVD:

- PRIMERGY BX900 S1 Blade Server System – Operating Manual
- PRIMERGY BX400 S1 Blade Server System – Operating Manual
- PRIMERGY Server Systems – RemoteDeploy
- PRIMERGY BX Blade Server Systems – LAN Connection Blade
- ServerView User Guide



For further information on updating BX components please refer to the Operating Manual of your BX system.

1.5 PRIMERGY BX900

The PRIMERGY BX900 Blade Server system is a modular server system that integrates up to 18 server modules, eight connection blade modules and two Management Modules (MMB).

The InfiniBand connection blade module provides networking and Switch functions to PRIMERGY BX900 Blade Server system. The Management Module offers a single point of control for the PRIMERGY BX900 Blade Server system.

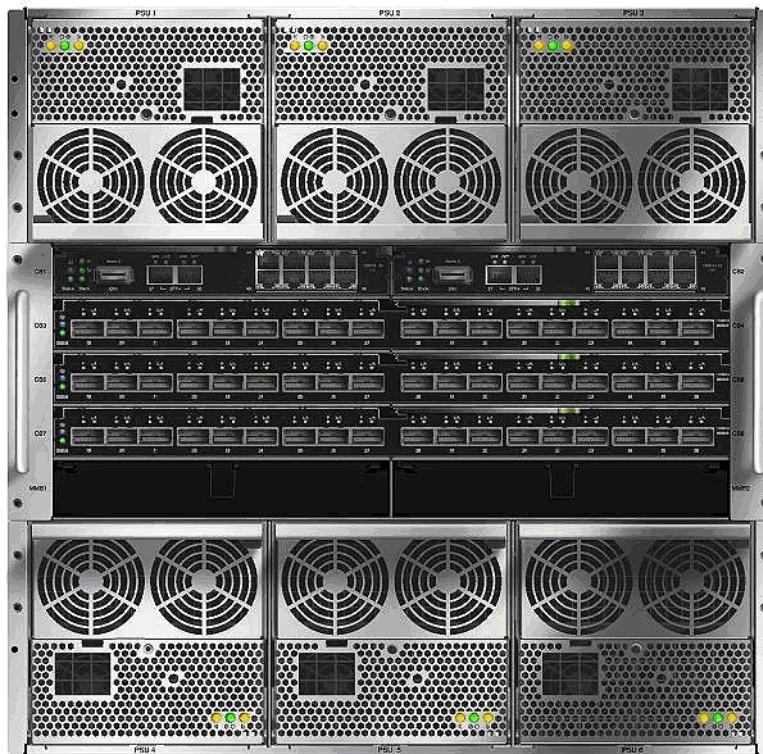


Figure 1: PRIMERGY BX900 Blade Server

1.6 Inserting Connection Blade into PRIMERGY BX900

In order to support 4x QDR InfiniBand, 4 lanes per one port are required in backplane. Due to this reason, the form factor of connection blade is double wide - single height.

The eject lever of the connection blade has a green touch button. The places at which all components may be touched to prevent them from being damaged when they are assembled or disassembled are also marked green.

Rear view of enclosure			
Fabric 1	CB1	CB2	
Fabric 2	CB3	CB4	
Fabric 3	CB5	CB6	
Fabric 4	CB7	CB8	

Figure 2: Rear view of enclosure

1.7 Connecting Mezzanine Cards on Server Blade

The figure below indicates the connection relation between connection blade (CB) and Mezzanine Card on server blade. InfiniBand (IB) Mezzanine Cards can be set to both Mezzanine I slot and Mezzanine II slot. When Mezzanine Card is installed in Mezzanine II slot, the dual port of the card will be available and one of two ports will be connected to CB in CB5/6, the other one will be CB7/8. While installed in Mezzanine slot I, only one port will be utilized and connected to the CB in CB3/4. When the two connection blades are installed in CB5/6 and CB 7/8, other types of connection blades like SB11/SB11a can be used in CB1/2/3/4 and, SBAX2 and Brocade FC connection blade in CB3/4 in parallel with IB connection blade.

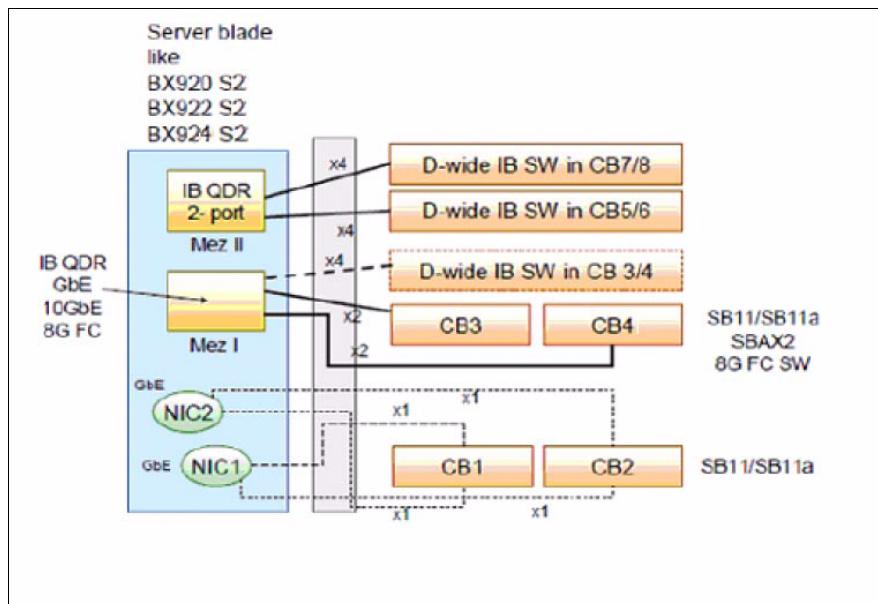


Figure 3: Connection relation between server blade and connection blade

1.8 Description of InfiniBand Connection Blade

The PRIMERGY BX900 InfiniBand connection blade module features 18 internal and 18 external 4x10Gbps (QDR) ports. The below figure shows the front view with the 18 external ports, whereas the internal ones are connected to the server blades over Midplane.



Figure 4: InfiniBand connection blade

1.8.1 Cables

For best performance use the InfiniBand connection blade with QSFP (Quad Small Form factor Pluggable) connectors. Copper and optical variants can be used.

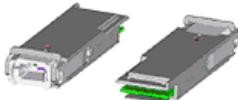


Figure 5: QSFP connectors

1.8.2 LEDs

Connection Blade LEDs

To indicate overall status of the connection blade, two LEDs are located at the left side of the front panel.

LED	Position	Color	Function
LED-A (Identify LED)	Left side: top	Blue	Identify active
		Off	Identify inactive
LED-B (Health LED)	Left side: bottom	Green	Power on
		Amber blinking	Error
		Off	Power off

Table 1: Connection blade LEDs

Port LEDs

For every external port a pair of LEDs exists.

LED	Position	Color	Function
LED-A (Port Physical Link)	Above port: right	Green	Link up
		Off	No Link
LED-B (Port Logical Link)	Above port: left	Yellow	Link up
		Yellow blinking	Link active
		Off	Link down

Table 2: Port LEDs

1.9 Technical Data

1.9.1 IB Connection Blade

Internal Ports:	18
External Ports:	18
Lanes per Port:	4 (4x)
Max Speed per Lane:	10Gbps (QDR)

Product Weight:	9.4 Lbs / 4.3 kg
Form Factor:	Double wide
Height:	27.9 mm / 1.1"
Width:	388.2 mm / 15.3"
Depth:	267.9 mm / 10.6"

1.9.2 IB Mezzanine Card

Form factor:	BX900 Standard form factor
Server blades supported:	BX920 S2
	BX922 S2
	BX924 S2
	BX92 S3
	BX924 S3
Max. number of Mezzanine Cards per blade:	2
Host interface:	x8 PCI-Express Gen2
Interface to midplane:	2 * 4x10Gbps (QDR)

(Consider restrictions mentioned in section “[Inserting Connection Blade into PRIMERGY BX900](#)” on page 9 and section “[Connecting Mezzanine Cards on Server Blade](#)” on page 10).

2 Linux Installation

2.1 Overview

This chapter describes how to install a single host machine with Fujitsu InfiniBand hardware installed. A blade can be properly installed with all required InfiniBand drivers and software during Red Hat Enterprise Linux installation.

If you purchased the products from Fujitsu Japan, you can download it from <http://jp.fujitsu.com/platform/server/primergy/downloads/>.



Please note, that only **Red Hat Enterprise Linux Version 5.4 or higher** is supported!

The chapter includes the following sections:

- “Identifying the GUID” on page 35
- “Native InfiniBand Installation in RHEL” on page 17
- “OFED Installation by means of SVIM” on page 20
- “Updating Firmware after Installation” on page 27
- “Subnet Manager” on page 27
- “Network Configuration Information” on page 28

2.2 Software Requirements

Required Disk Space for Linux Installation

400 MB

Operating System

Linux operating system (x86_64 only)

Installer Privileges

The installation requires administrator privileges on the target machine.

i To reach best performance, the *Enhanced Idle Power State* option in the BIOS setup *Advanced* menu must be disabled.

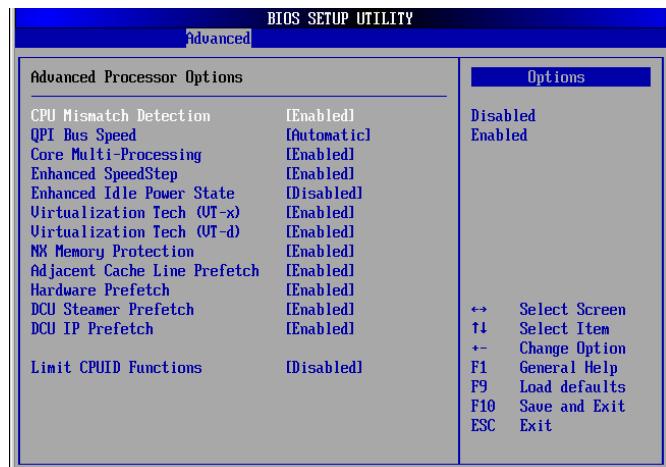


Figure 6: BIOS Setup Utility

2.3 Native InfiniBand Installation in RHEL

This chapter describes the installation method of OFED for Red Hat Enterprise Linux Version 5.4.

2.3.1 Setup InfiniBand during Linux Installation

All required InfiniBand drivers and tools can be set up during the installation of Red Hat Enterprise Linux.

Step 1

Run the installation routine until the "Package Installation Defaults" screen appears. This screen varies depending on the version of Red Hat Enterprise Linux you are installing.



Figure 7: Red Hat: package installation defaults

Step 2

Select the "Customize now" option on the screen. The next screen named "Package Group Selection" shows a list of groups containing additional packages. Selecting the list item "Base Systems" displays the detailed view on the right side.



Figure 8: Red Hat: package group selection

Step 3

To install OFED, select the "Administrator Tools", and press the "Optional packages" button while the OFED list item is selected and mark the "All packages" entry in the new window.

Step 4

Continue the installation process as normal.

2.3.2 Loading InfiniBand Drivers at System Startup

The complete OFED driver stack can be loaded automatically at system startup. Alternatively only the "Core" modules will run.

The *openibd* script residing in */etc/init.d/* is responsible for this task. It can be configured by editing the */etc/infiniband/openib.conf*.

To execute this script at startup, it is added as a service at the appropriate run level and activated.

```
host1# chkconfig --add openibd
host1# chkconfig openibd on
```

After rebooting the system, all important InfiniBand modules will be loaded into the kernel.

2.3.3 Installation Results

Kernel modules

The kernel modules are installed under:

- InfiniBand subsystem: `/lib/modules/`uname -r`/kernel/drivers/infiniband/`
- mlx4 driver: Under `/lib/modules/`uname -r`/kernel/drivers/net/mlx4` you will find `mlx4_core.ko`, `mlx4_en.ko` (and `mlx4_fc` if you ran the installation with `--with-fc`)
- RDS: `/lib/modules/`uname -r`/kernel/net/rds/rds.ko`
- Bonding module: `/lib/modules/`uname -r`/kernel/drivers/net/bonding/bonding.ko`

Startup script

The script `openibd` is installed under `/etc/init.d/`. This script can be used to load and unload the software stack.

Configurations files

The `/etc/ofed` directory is created with the file `openib.conf`. The `openib.conf` file contains the list of modules that are loaded when the `openibd` script is used.

If OpenSM is installed, the daemon `opensmd` is installed under `/etc/init.d/` and `opensm.conf` is installed under `/etc/ofed`.

Other files

The `90-ib.rules` file is installed under `/etc/udev/rules.d/`.

Man pages will be installed under `/usr/share/man/`.

2.3.4 Post-installation Notes

User-space Memory Limitations

The script adds the following lines to `/etc/security/limits.conf` for the user space components such as MPI:

- * soft memlock unlimited
- * hard memlock unlimited

These settings make the amount of memory that can be pinned by a user space application unlimited. If desired, tune the value unlimited to a specific amount of RAM.

Subnet Manager required

For your machine to be part of the InfiniBand fabric, a Subnet Manager must be running on one of the fabric nodes. At this point, Mellanox OFED for Linux has already installed the OpenSM Subnet Manager on your machine. For details on starting OpenSM, see “[Subnet Manager](#)” on page 27.

IPoIB function

The IPoIB function shouldn't be set to the IB port which is the port1 on the slot1. The limitation is applied to RHEL5 and RHEL6.

IPoIB adapters parameter limitation

In case the IPoIB adapters are set by using ifcfg-ib* files, the parameter 'NM_CONTROLLED' should be set to 'no'. The limitation is applied only to RHEL6.

2.4 OFED Installation by means of SVIM

The installation of OFED gets decided on the package selection in ServerView Installation manager (SVIM).

- When all packages were selected, OFED is installed within OS installation.
- When the default package groups were selected, OFED is not installed.

Please install OFED by the following steps after the OS installation.

2.4.1 In the event of RHEL5.4

This chapter describes the installation of OFED for Red Hat Enterprise Linux Version 5.4.

1. Insert OS DVD in the DVD drive, and make sure of the mount of DVD.

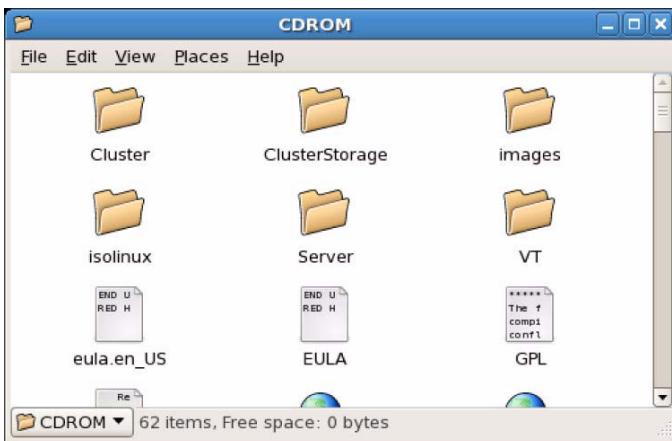


Figure 9: OS DVD

It is assumed that the mount position of the DVD is `/media/CDROM` in the following procedures.

2. Check the directories in the OS DVD which include the *rpm* files.
In the following example, the *rpm* files are included in *Cluster*, *ClusterStorage*, *Server* and *VT* directories.
3. Run Package Manager and click *Edit - Repositories*.

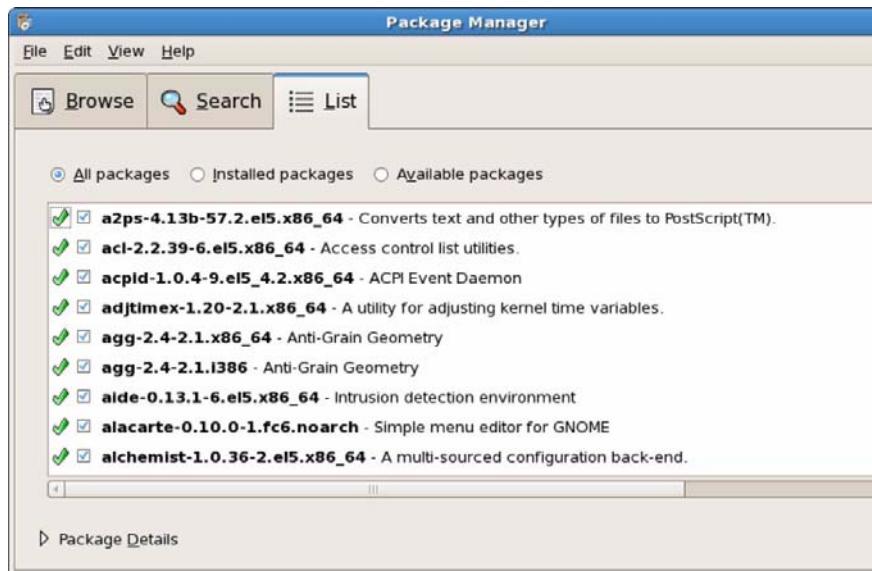


Figure 10: Package Manager

4. Select *Add* to add the repository.

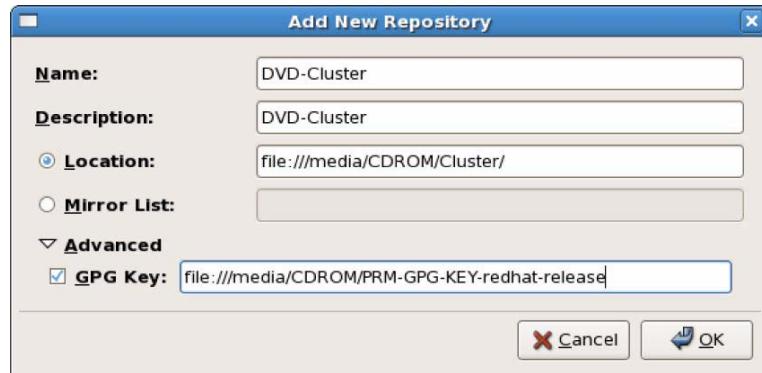


Figure 11: Add new Repository

Item	Name
Name	Arbitrary name
Description	Arbitrary explanation
Location (L)	<p>The format is as follows: <code>file:///< mount position of DVD media>/<directory name here the inclusion of rpm file></code></p> <p>Here, the mount position of DVD media is assumed to be <code>/media/CDROM</code> .</p> <p>The directory names where the <i>rpm</i> files are included are <i>Cluster</i>, <i>ClusterStorage</i>, <i>Server</i>, and <i>VT</i> in Red Hat Enterprise Linux Version 5.4.</p>
GPG Key (G)	Path to GPG key file.

In the following example, GPG key file *RPM-GPG-KEY-redhat-release* is assumed to be included in the OS DVD.

5. Make sure that *RPM-GPG-KEY-redhat-release* is included in the OS DVD.
6. Enter the path to the file in *GPG Key*.
7. Set the repository for all the directories.
8. Make sure that the directories created above are ticked and other repositories are not ticked in repository manager.

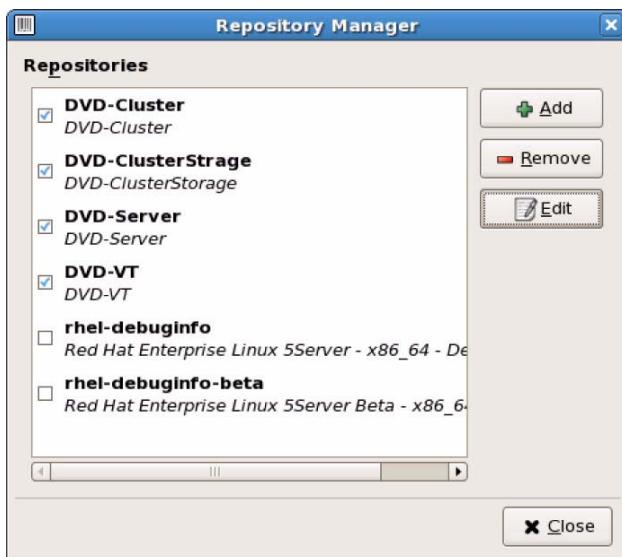


Figure 12: Repository Manager

9. Close the Repository Manager and the Package Manager.

10. Run the terminal and execute the following command.

```
# yum repolist
```

11. Run the Package Manager.

12. Tick *OpenFabrics Enterprise Distribution* group in *Base System*.

13. Tick all the packages in the group.

14. Click *Apply*.

OFED is installed.

2.4.2 In the event of RHEL6.0

This chapter describes the installation of OFED for Red Hat Enterprise Linux Version 6.0.

Please install OFED by means of OS DVD after the OS installation.

1. Insert OS DVD in the DVD drive, and make sure that the DVD is mounted.

2. Add the following statement to `/etc/yum.conf`.
`group_package_types=mandatory,default,optional`
3. Make the `.repo` file, then save the file to `/etc/yum.repos.d/`
In the following examples, DVD media is mounted to
`/media/RHEL_6.0_x86_64 Disc 1`, the repository name is `dvd`.

Example :

```
/etc/yum.repos.d/.repo
[dvd]
baseurl=file:///media/RHEL_6.0\ x86_64\ Disc\ 1
enabled=1
gpgcheck=1
gpgkey=file:///media/RHEL_6.0\ x86_64\ Disc\ 1/RPM-GPG-KEY-
redhat-release
```

4. Execute the following command on the terminal: (the repository name is `dvd`.)

```
[root@ localhost ]# yum clean all
[root@ localhost ]# yum --disablerepo=\* --enablerepo=dvd
makecache
[root@ localhost ]# yum --disablerepo=\* --enablerepo=dvd
groupinstall Infiniband -y
```

OFED is installed.

2.4.3 Post-installation Notes

This section describes how to set `mlx4` parameters.

Add a line to a file:

- If you set to RHEL6.1 + MellanoxOFED, please add the line to the `/etc/modprobe.d/mlx4.conf`.
- If there is no `mlx4.conf` file, please create it.
- How to set `mlx4` parameters:
`options mlx4_core parameter=<value>`

How to change MR

log_num_mtt

Log maximum number of memory translation table segments per HCA
(default is 20; max is 20)

log_mtts_per_seg

Log number of MTT entries per segment (default is 3; max is 7).

The following table shows the recommended values of *log_num_mtt* and *log_num_per_seg*. (The values change with the main memory capacity.)

main memory capacity	16 GB	32 GB	64 GB
<i>log_num_mtt</i>	18	19	20
<i>log_mtts_per_seg</i>	4	4	4

After the parameter has been changed, you will need to reboot.

How to change QP number of HCA

log_num_qp

Log maximum number of QP number per HCA (default is 18).

Set value is the value of the powers of two.

Example: 18 (default) = 26,2144 (256K), 21 = 209,7152 (2M)

Example: */etc/modprobe.d/mlx4.conf*

options mlx4_core log_num_mtt=20 log_mtts_per_seg=4 log_num_qp=18

After the parameter has been changed, you will need to reboot.

IPoIB adapters parameter limitation

If you create (or edit) the *ifcfg-ib** files to configure an IPoIB, the parameter *NM_CONTROLLED* should be set to *no*.

The limitation is applied to RHEL6 only.

Example: */etc/sysconfig/network-scripts/ifcfg-ib0*
DEVICE="ib0"
IPADDR="192.168.210.100"
NETMASK="255.255.255.0"

```
BROADCAST=192.168.210.255
NM_CONTROLLED="no"
NOBOOT="yes"
```

Since there is a problem peculiar to RHEL6.1, communication by IPoIB may not be able to be performed. Please add the following parameters:

PREFIX=<value>

Set to the value corresponding to the *NETMASK*.

Example:

NETMASK=255.255.255.0

PREFIX=24

2.5 Updating Firmware after Installation

If you wish to burn newer firmware, you have to download it from Fujitsu Technologies' Web site (<http://support.ts.fujitsu.com/com/support/downloads.html> > *Driver & Downloads*).

If you purchased the products from Fujitsu Japan, you download it from <http://jp.fujitsu.com/platform/server/primergy/downloads/>.

2.6 Subnet Manager

At least one Subnet Manager(SM) is present on each subnet. Each SM resides on a port of IB Mezzanine card or IB connection blade. When there are multiple SMs on a subnet, the master SM is one, and the remaining SMs become standby SMs. The master SM has a role to initialize and configure an Infiniband subnet.

OpenSM is an Infiniband compliant SM. The following sections describe how to use OpenSM.

2.6.1 OpenSM

OpenSM is one Subnet Manager distributed with RedHat.

This chapter describes how to configure which server becomes the master SM. Please set up the server which will be the master using the following reference. Regarding the other settings of OpenSM, please refer to the command help of *opensmd*.

Basically, a server with the lowest GUID will be the master SM within a subnet. If you intend to choose the master SM, you can configure the priority which range is from 0 to 15. 0 is the lowest and 15 is the highest.

Example for RHEL5:

```
# /etc/init.d/opensmd start -p 15
```

2.6.2 Confirmation of OpenSM running

Enter the following command to confirm that OpenSM is running:

```
# /etc/init.d/opensmd status (RHEL5)
# /etc/init.d/opensm status (RHEL6)
```

OpenSM is running properly, if the result of this command displays "running".

Enter the following command to run OpenSM, if the result displays "stopped":

```
# /etc/init.d/opensmd start (RHEL5)
# /etc/init.d/opensm start (RHEL6)
```

2.6.3 Confirmation of the master SM

Enter the *ibstat* command on the system server, then "Base LID" and "SM LID" are displayed.

The IB Mezzanine card or IB connection blade on which these two LIDs are the same is the master SM.

2.7 Network Configuration Information

The following information (1)-(6) is needed to maintain and troubleshoot an InfiniBand network.

The information should be saved when configuring the InfiniBand network.

1. Network configuration diagram

2. Configuration sheet
3. The server on which the MASTER is found
4. The server on which the STANDBY is found
5. The result of the *ibnetdiscover* command
6. GUID list and arrangement drawing of device location plan

2.7.1 Network configuration diagram

Example:

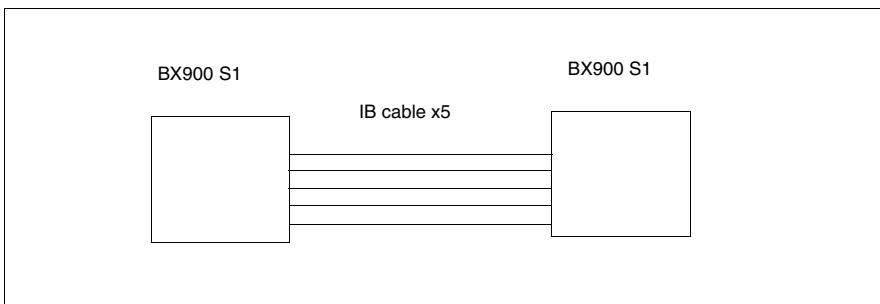


Figure 13: Network configuration diagram

2.7.2 Configuration sheet

Server blade	mezz# 1 GUID	mezz# 2 GUID	OS	host-name	Static /DHCP	iRMC	eth0	eth1	eth2	eth3	ib0	ib1	ib2	ib3
Slot1														
Slot2														
Slot3														
Slot4														
Slot5														
Slot6														
Slot7														
Slot8														

Table 3: Configuration sheet

Slot9											
Slot10											
Slot11											
Slot12											
Slot13											
Slot14											
Slot15											
Slot16											
Slot17											
Slot18											
	GUID		host-name	Static /DHCP		eth0	ib0	ntp	snmp		
CB3											
CB5											
MMB1-Agent											
MMB1-Service											

Table 3: Configuration sheet

IP Address= . . .XX

Root password:

How to fill in:

- Please fill in the GUID of the Mezzanine Card installed in Mezzanine I slot in the column of "mezz#1 GUID". Similarly, please fill in the GUID of the Mezzanine Card installed in Mezzanine II slot in the column of "mezz#2 GUID".
- Please fill in the host name in the column of "hostname" when you assign the host name to each server blade.
- Please fill in the IP addresses(XX) in the column of "iRMC", "eth0", "eth1", "eth2", "eth3", "ib0", "ib1", "ib2", and "ib3" when you use the static IP.

IP Address = 192.168.11.XX

role of description

mezz#1 GUID	mezz#2 GUID	OS	Hostname	Static /DHCP	iRMC	eth0	eth1	eth2	eth3	ib0	ib1	ib2	ib3
0002c9030003bf7a	0002c9030003bf6	RHEL5.4	bx920-02-01	static	101	102	103	104	105	-	-	-	-
0002c903000413c0	0002c90300041514	RHEL5.4	bx920-02-02	static	106	107	108	109	110	-	-	-	-
GUID			hostname	Static /DHCP		eth0	ib0	ntp	snmp				
				static		33	-	-	-				
				static		34	-	-	-				
agent													
device													

Figure 14: Example of configuration sheet

2.7.3 The result of the command "ibnetdiscover"

GUID list

It is useful to provide the text file (:.txt) which associates a GUID with the name of the device so that the *ibnetdiscover* command displays the result more meaningfully.

format

GUID(hex number) "Hostname and Mezzanine slot written in the Configuration sheet"

```
# bx900-02 CARD
0x0002c903000413c0 "bx900-02-01 HCA-1"
0x0002c90300041514 "bx900-02-01 HCA-2"
0x0002c9030003bf7a "bx900-02-02 HCA-1"
0x0002c9030003bf6a "bx900-02-02 HCA-2"
0x0002c903000413b8 "bx900-02-03 HCA-1"
0x0002c90300041488 "bx900-02-03 HCA-2"
0x0002c90300041408 "bx900-02-04 HCA-1"
0x0002c903000412f0 "bx900-02-04 HCA-2"
0x0002c90300041414 "bx900-02-05 HCA-1"
0x0002c90300041348 "bx900-02-05 HCA-2"
0x0002c903000412ec "bx900-02-06 HCA-1"
0x0002c9030004139c "bx900-02-06 HCA-2"
0x0002c9030003bf9e "bx900-02-07 HCA-1"
0x0002c9030004141c "bx900-02-07 HCA-2"
0x0002c9030003bf86 "bx900-02-08 HCA-1"
0x0002c9030003bfca "bx900-02-08 HCA-2"
0x0002c9030003bfa6 "bx900-02-09 HCA-1"
0x0002c90300041314 "bx900-02-09 HCA-2"
0x0002c903000413e8 "bx900-02-10 HCA-1"
0x0002c903000413f0 "bx900-02-10 HCA-2"
0x0002c9030003bfba "bx900-02-11 HCA-1"
0x0002c9030003bf7e "bx900-02-11 HCA-2"
0x0002c9030003bf92 "bx900-02-12 HCA-1"
0x0002c9030003bfa2 "bx900-02-12 HCA-2"
0x0002c9030003d138 "bx900-02-13 HCA-1"
0x0002c90300041304 "bx900-02-13 HCA-2"
0x0002c9030003bf96 "bx900-02-14 HCA-1"
0x0002c9030003bf72 "bx900-02-14 HCA-2"
0x0002c90300041460 "bx900-02-15 HCA-1"
```

```
0x0002c90300041f48 "bx900-02-15 HCA-2"  
0x0002c903000412fc "bx900-02-16 HCA-1"  
0x0002c903000412f4 "bx900-02-16 HCA-2"  
0x0002c903000413fc "bx900-02-17 HCA-1"  
0x0002c90300041f38 "bx900-02-17 HCA-2"  
0x0002c9030003bf82 "bx900-02-18 HCA-1"  
0x0002c9030003bf8e "bx900-02-18 HCA-2"
```

```
# BX900-02 IB SWITCH  
0x0002c9020040c7a8 "bx900-02-IB-SW-CB3"
```

```
# BX900-04 IB SWITCH  
0x0002c9020040c790 "bx900-04-IB-SW-CB3"
```

```
# bx900-04 CARD  
0x0002c903000413ec "bx900-04-01 HCA-1"  
0x0002c903000447b0 "bx900-04-02 HCA-1"
```

"ibnetdiscover" command execution procedure

1. Connect to the CLI of IB connection blade via MMB (MMB Console Redirection).
2. Execute the command *swinfo*

```
<BX900S1-CB3> swinfo  
<BX900S1-CB3 Information:  
  GUID: 0x0002c9020040c7a8  
  MAC: 00:02:c9:11:20:5a  
  IS4 Firmware revision: 7.2.326  
  CPU Firmware revision: Mellanox release mlnx405ex-1.0.2 build 2009-07-19
```
3. In the example above the GUID of IB connection blade is "0x0002c9020040c7a8"
4. Execute the command *ibnetdiscover* on the server blade of slot1.

```
[root@bx900-02-01 ~]# ibnetdiscover --node-name-map /root/mmap.txt
#
# Topology file: generated on Sat Aug 1 01:27:33 2009
#
# Max of 3 hops discovered
# Initiated from node 0002c903000413c0 port 0002c903000413c2

vendid=0x2c9
devid=0xb36
sysimguid=0x2c9020040c7a8
switchguid=0x2c9020040c7a8(2c9020040c7a8)

Switch 36 "S-0002c9020040c7a8" (a) # "bx900-02-IB-SW-CB3" enhanced port 0 lid 14 lmc 0
[23] C "S-0002c9020040c790" [23] # "bx900-04-IB-SW-CB3" lid 9 4xQDR
[22] "S-0002c9020040c790" [22] # "bx900-04-IB-SW-CB3" lid 9 4xQDR
[21] "S-0002c9020040c790" [21] # "bx900-04-IB-SW-CB3" lid 9 4xQDR
[20] "S-0002c9020040c790" [20] # "bx900-04-IB-SW-CB3" lid 9 4xQDR
[19] "S-0002c9020040c790" [19] # "bx900-04-IB-SW-CB3" lid 9 4xQDR
[18] "H-0002c9030003bf82" [2](2c9030003bf84) # "bx900-02-18 HCA-1" lid 7 4xQDR
[17] "H-0002c903000413fc" [2](2c903000413fe) # "bx900-02-17 HCA-1" lid 22 4xQDR
[16] "H-0002c903000412fc" [2](2c903000412fe) # "bx900-02-16 HCA-1" lid 21 4xQDR
[15] "H-0002c90300041460" [2](2c90300041462) # "bx900-02-15 HCA-1" lid 5 4xQDR
[14] "H-0002c9030003bf96" [2](2c9030003bf98) # "bx900-02-14 HCA-1" lid 11 4xQDR
[13] "H-0002c9030003d138" [2](2c9030003d13a) # "bx900-02-13 HCA-1" lid 4 4xQDR
[12] "H-0002c9030003bf92" [2](2c9030003bf94) # "bx900-02-12 HCA-1" lid 10 4xQDR
[11] "H-0002c9030003bfba" [2](2c9030003bfbc) # "bx900-02-11 HCA-1" lid 17 4xQDR
[10] "H-0002c903000413e8" [2](2c903000413ea) # "bx900-02-10 HCA-1" lid 18 4xQDR
[9] "H-0002c9030003bfa6" [2](2c9030003bfa8) # "bx900-02-09 HCA-1" lid 13 4xQDR
[8] "H-0002c9030003bf86" [2](2c9030003bf88) # "bx900-02-08 HCA-1" lid 8 4xQDR
[7] "H-0002c9030003bf9e" [2](2c9030003bf9a0) # "bx900-02-07 HCA-1" lid 12 4xQDR
[6] "H-0002c903000412ec" [2](2c903000412ee) # "bx900-02-06 HCA-1" lid 19 4xQDR
[5] "H-0002c90300041414" [2](2c90300041416) # "bx900-02-05 HCA-1" lid 3 4xQDR
[4] "H-0002c90300041408" [2](2c9030004140a) # "bx900-02-04 HCA-2" lid 2 4xQDR
[3] "H-0002c903000413b8" [2](2c903000413ba) # "bx900-02-03 HCA-1" lid 16 4xQDR
[2] "H-0002c9030003bf7a" [2](2c9030003bf7c) # "bx900-02-02 HCA-1" lid 6 4xQDR
[1] "H-0002c903000413c0" [2](2c903000413c2) # "bx900-02-01 HCA-1" lid 1 4xQDR

vendid=0x2c9
devid=0xb36
sysimguid=0x2c9020040c790(2c9020040c790)
switchguid=0x2c9020040c790(2c9020040c790)

Switch 36 "S-0002c9020040c7a8" # "bx900-02-IB-SW-CB3" enhanced port 0 lid 9 lmc 0
[2] "H-0002c903000447b0" [2](2c903000447b2) # "bx900-04-02 HCA-1" lid 15 4xQDR
[1] "H-0002c903000413ec" [2](2c903000413ee) # "bx900-04-01 HCA-1" lid 20 4xQDR
[23] "S-0002c9020040c7a8" [23] # "bx900-02-IB-SW-CB3" lid 14 4xQDR
[22] "S-0002c9020040c7a8" [22] # "bx900-02-IB-SW-CB3" lid 14 4xQDR
[21] "S-0002c9020040c7a8" [21] # "bx900-02-IB-SW-CB3" lid 14 4xQDR
[20] "S-0002c9020040c7a8" [20] # "bx900-02-IB-SW-CB3" lid 14 4xQDR
[19] "S-0002c9020040c7a8" [19] # "bx900-02-IB-SW-CB3" lid 14 4xQDR
```

Figure 15: *ibnetdiscover* command: GUID list and arrangement drawing of devices

ibnetdiscover format

1. Provide the GUID list file (mmap.txt) in advance.
2. Search for the GUID obtained by the command `swinfo` in the result of `ibnetdiscover` ("S-" indicates IB connection blade, see *a*)
3. There is a string "lid 14" in the same line as the GUID referred above. 14 indicates the LID of IB connection blade (see *b*).
4. The leftmost figure in each line "[23] to [1]" indicate the port number which is connected to a device.
The maximum number of the port is 36. Thus if all ports are connected, "[1] to [36]" is displayed.
In this example, there are connections as follows (see *c*).
5. "4x QDR" indicates the Quad Data Rate connection.
In case this value is "SDR" or "DDR" or "QDR", there is something wrong with the cable connection (see *d*).

GUID list of IB connection blade, the layout chart, and rack loading list is useful for the maintenance.

2.8 Identifying the GUID

The IB Mezzanine card has 4 different GUIDs (Globally Unique Identifier): Node GUID, System GUID, GUID for Port1 and GUID for Port2). The Node GUID is listed on the label on the back of the product.



Figure 16: Label on the back of the IB Mezzanine Card

How to identify the node GUID:

- ▶ Log in to the server blade as root.
- ▶ Run the open terminal or telnet/ssh.
- ▶ Execute the *ibstat* command, and identify the node GUID ("Node GUID: 0x*****"). The *ibstat* command is available after installation of OFED Stack for Linux.

The node GUID is "0002c903000413b8" in the table below.

```
[root@bx900-02-03 ~]# ibstat
CA 'mlx4_0'
CA type: MT26428
Number of ports: 2
Firmware version: 2.6.922
Hardware version: a0
Node GUID: 0x0002c903000413b8
System image GUID: 0x0002c903000413bb
Port 1:
State: Down
Physical state: Polling
Rate: 10
Base lid: 0
LMC: 0
SM lid: 0
Capability mask: 0x02510868
Port GUID: 0x0002c903000413b9
Port 2:
State: Active
Physical state: LinkUp
Rate: 40
Base lid: 71
LMC: 0
SM lid: 1
Capability mask: 0x02510868
Port GUID: 0x0002c903000413ba
CA 'mlx4_1'
CA type: MT26428
Number of ports: 2
Firmware version: 2.6.922
Hardware version: a0
Node GUID: 0x0002c903000446f0
System image GUID: 0x0002c903000446f3
Port 1:
State: Down
Physical state: Polling
Rate: 10
Base lid: 0
LMC: 0
SM lid: 0
Capability mask: 0x02510868
Port GUID: 0x0002c903000446f1
Port 2:
State: Active
Physical state: LinkUp
Rate: 40
Base lid: 91
LMC: 0
SM lid: 1
Capability mask: 0x02510868
Port GUID: 0x0002c903000446f2
[root@bx900-02-03 ~]#
```

3 Windows Installation

3.1 Overview

This chapter describes how to install a single host machine with Fujitsu InfiniBand hardware installed. A server can be properly installed with all required InfiniBand drivers and software during Windows (HPC) Server 2008 R2 installation or afterwards by using the Win OpenFabrics installation or OpenSM setup.



Please note that only **Windows Server 2008 R2** and **Windows HPC Server 2008 R2** are supported!

3.1.1 Software Requirements

Required Disk Space for Windows Installation

100 MB

Operating System

Windows (HPC) Server 2008 R2

Installer Privileges

Installation requires administrator privileges on the target machine.

i To reach best performance, the *Enhanced Idle Power State* option in the BIOS setup *Advanced* menu must be disabled.

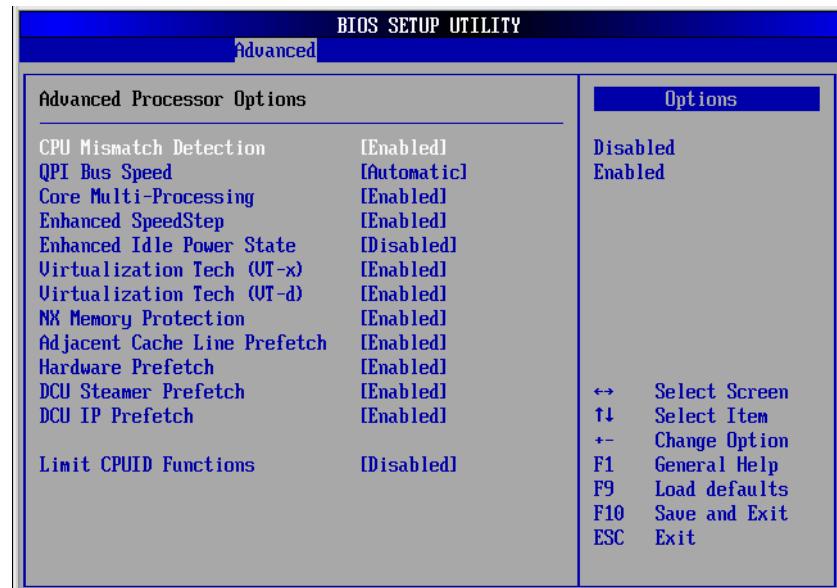


Figure 17: BIOS Setup Utility

3.2 Installation of WinOF

When WinOF was not installed by Server View Installation Manager (SVIM), please follow the following steps. If WinOF was installed, please jump to step 10.

1. Make sure that the InfiniBand device is shown in the Device Manager.



Figure 18: Display of InfiniBand devices before installation of WinOF

2. Log on to Windows as administrator, and then double-click the file of WinOF in ServerView Suite DVD1:

D:\DRV\InfiniBand\Mellanox\MT26428\x64W2K8R2\

3. Click [Next>] after the following window appears.



Figure 19: InstallShield Wizard

4. After the following window appears, read the License Agreement carefully. If you accept, click "I accept the terms in the license agreement" and then click [Next>].

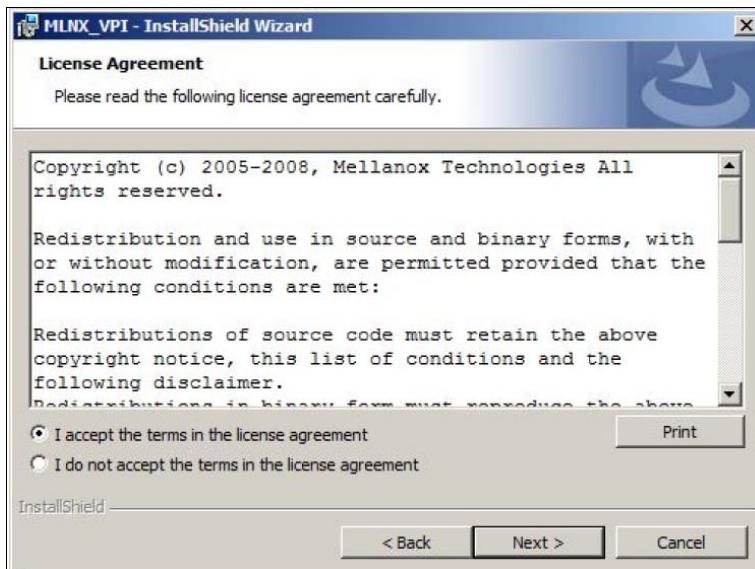


Figure 20: License Agreement

5. Click [Next>] after the following window appears (if you want to change the installation folder, specify the folder you want to install on by clicking [Change...]).



Figure 21: Destination Folder

6. After the following window appears, remove the check mark from "Check this box to configure your system for maximum 10GigE performance (Recommended)", and then click [Next>].

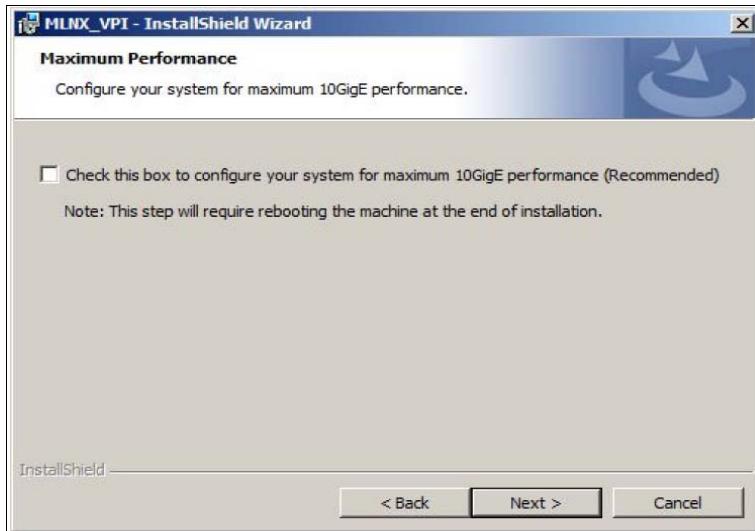


Figure 22: Maximum performance

7. Click [Install] after the following window appears.

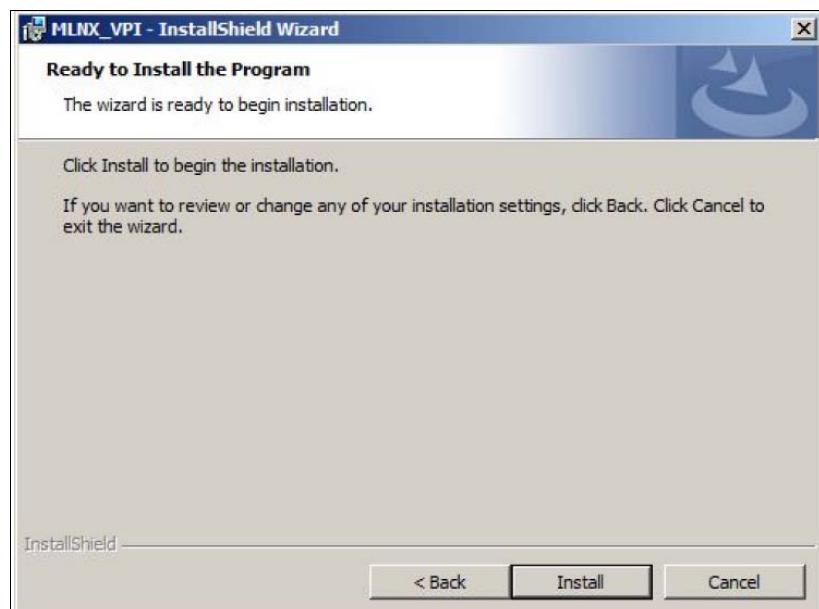


Figure 23: Installation

8. The following window appears after the installation process is finished. Click [Finish] without checking any boxes.



Figure 24: Installation completed

9. Make sure that InfiniBand devices are recognized properly in the device manager.

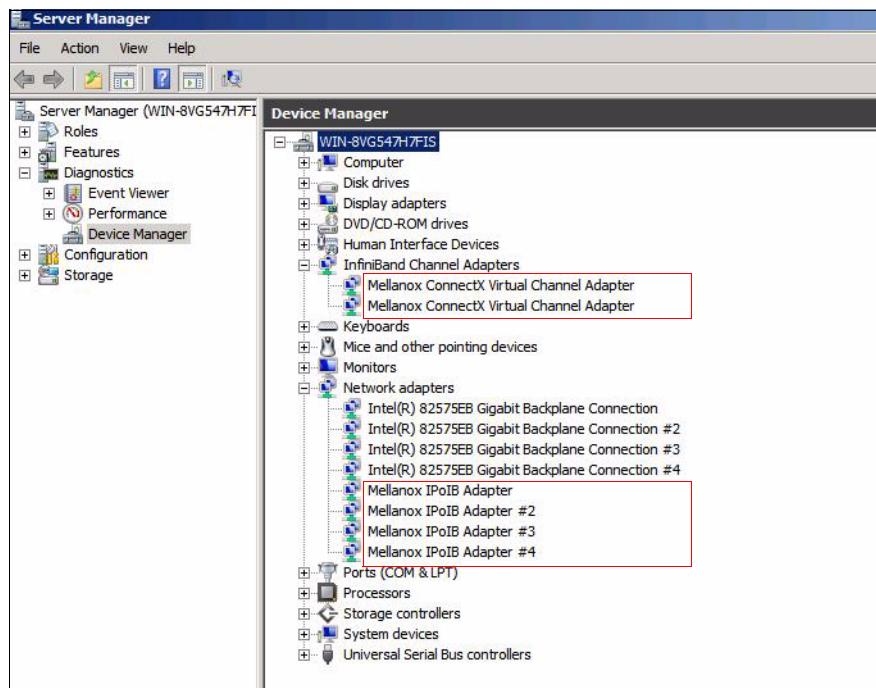


Figure 25: InfiniBand devices

10. After WinOF installation, OpenSM should be set up.

OpenSM is an InfiniBand compliant Subnet Manager (SM). At least one OpenSM is present on the InfiniBand network. OpenSM has a role to initialize and configure an InfiniBand network.

If HPC Cluster will consist of up to 16 nodes or less, OpenSM can run on the head node computer.

The 'OpenSM' service is registered to the **[Services]** Window after WinOF installation.

Double-click 'OpenSM' to open properties.

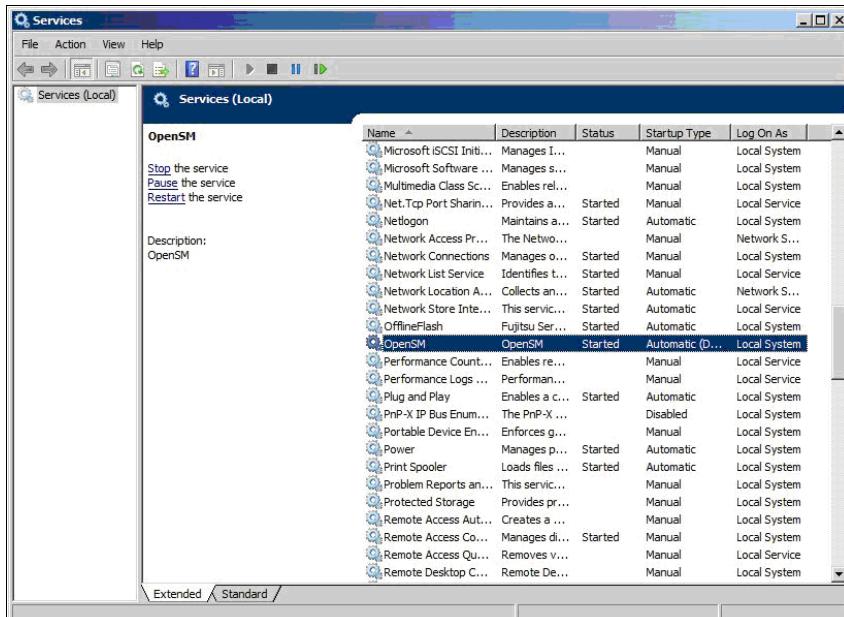


Figure 26: Configuration of OpenSM

11. Change 'Startup type' from *Manual* to *Automatic (Delay Start)*

12. Click the 'Start' button to start the OpenSM service.

13. Open the 'Network Connections' Window to check if IPoIB networks are linked up.

Disable the unused IPoIB adapters.

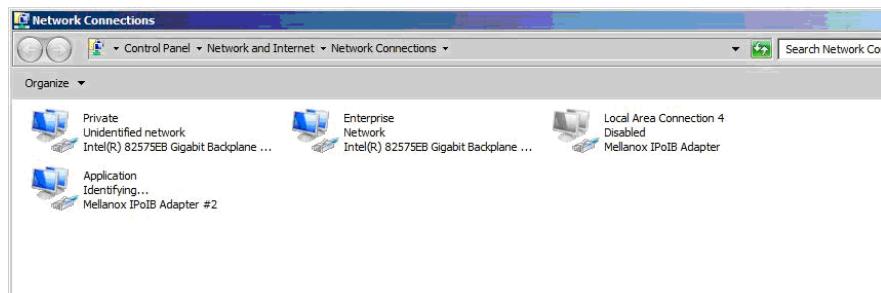


Figure 27: Network status

14. Since no IP address is assigned to this new IPoIB adapter, assign an IP address to it.

15. Run the command prompt as the Administrator privilege and execute the following command to enable 'Network Direct'.

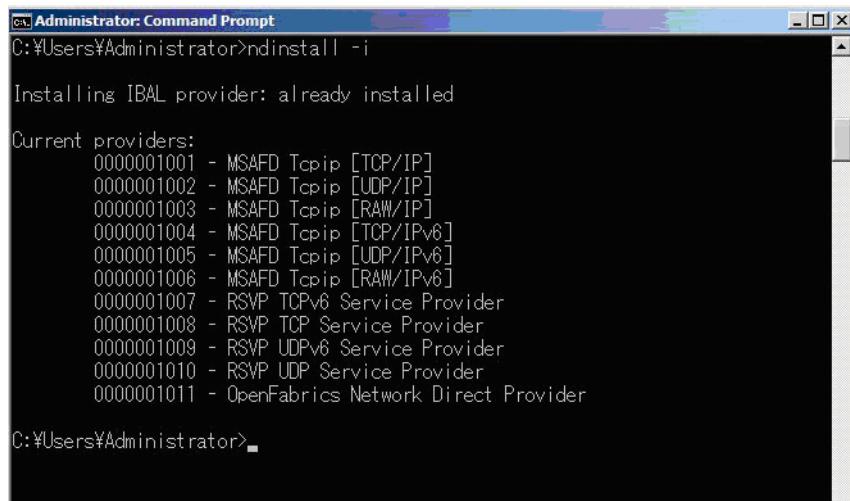
```
> ndinstall -i
```



Figure 28: Installation of Network Direct 1

16. In order to check if 'Network Direct' is enabled properly, execute the same command again.

The message 'Installing IBAL provider: already installed' is shown if it is enabled properly.



```
Administrator: Command Prompt
C:\Users\Administrator>ndinstall -i
Installing IBAL provider: already installed

Current providers:
0000001001 - MSAFD Tcpip [TCP/IP]
0000001002 - MSAFD Tcpip [UDP/IP]
0000001003 - MSAFD Tcpip [RAW/IP]
0000001004 - MSAFD Tcpip [TCP/IPv6]
0000001005 - MSAFD Tcpip [UDP/IPv6]
0000001006 - MSAFD Tcpip [RAW/IPv6]
0000001007 - RSVP TCPv6 Service Provider
0000001008 - RSVP TCP Service Provider
0000001009 - RSVP UDPv6 Service Provider
0000001010 - RSVP UDP Service Provider
0000001011 - OpenFabrics Network Direct Provider

C:\Users\Administrator>
```

Figure 29: Installation of Network Direct 2

4 IB Connection Blade Management

4.1 Connection

In this chapter the connection of IB connection blade via Management Blade is explained. Follow the procedures below.

Configurations for connecting to IB connection blade are explained in [section "Configuration" on page 54](#).

1. Login to Management Blade CLI via telnet.
Please refer to "Server View Management Blade" to see about the connection and the booting up of Management Blade CLI.
2. Select "(3) Console Redirection" and press the **[Enter]** key.
3. Select "(1) Console Redirect Connection Blade" and press the **[Enter]** key.

```
+-----+  
|       Console Redirection Table      page_3  
+-----+  
(1) Console Redirect Connection Blade  
(2) Set Return Hotkey , Ctrl+(a character) : Q  
(3) Set Console Redirection Timeout      : 900  
Enter selection or type (0) to quit:
```

Figure 30: Console redirection

4. Enter the slot number on which IB connection blade is installed and press the **[Enter]** key.
(Example : If IB connection blade is installed on slot3/4, enter "3" and press the **[Enter]** key)

```
Console Redirect Connection Blade page_3_1
+
(1) CB1:PY CB Eth Switch/IBP 1Gb 36/12 (Mode:Switch) - Stacking:N/A
(3) CB3:PY CB IB Switch 40Gb 18/18 (Mellanox) - Stacking:N/A
(5) CB5:PY CB IB Switch 40Gb 18/18 (Mellanox) - Stacking:N/A
Enter selection or type (0) to quit: 3
```

Figure 31: Console redirect connection blade

4.1.1 telnet connection

The user can execute the commands via telnet from the management console.

The method of connecting to IB connection blade via telnet is described below.

- ▶ Run a terminal software on the management console.
- ▶ Enter the command including the IP address referred to above "telnet 192.168.1.10" and press the **Enter** key.

The user connects to IB connection blade via telnet.

- ▶ Enter the username "maintenance" and the password "#yrotcaf" and press the **Enter** key.

```
[libuser@localhost ~]$ telnet 192.168.1.10
Trying 192.168.1.10...
Connected to 192.168.1.10 (192.168.1.10).
Escape character is '^]'.
Linux 2.6.27 (localhost) (14:53 on Monday, 24 August 2009)
BX900S1P00131-CB3 login: maintenance
Password:
```

4.1.2 Save and restore the configuration

The configuration data can be saved and restored by ftp command from PC.



Caution!

Look out for the following during the operation of save and restore:

- Do NOT turn off the power
- Check that there is no traffic between the nodes.
- Operate with NOT doing any setting from a management terminal.

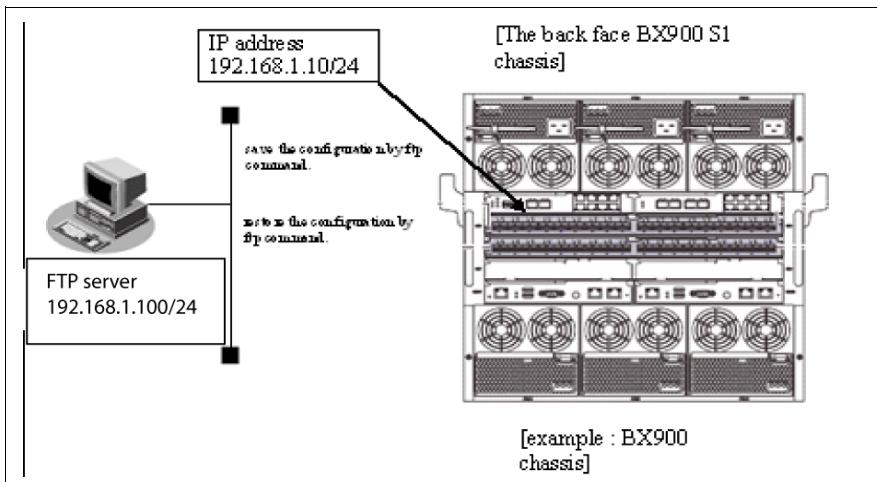


Figure 32: Save configuration

How to save the configuration

If you would like to save the configuration file as 'ibswb-conf' file for example, do the following:

1. Configure IP address, see [section “Configuration” on page 54](#).
2. Execute *saveconf* command from IB connection blade.
3. Copy *ibswb-conf* file to another PC or server by *ftp* or *scp* command, because under */usr/ftproot* files will disappear after IB connection blade reboots.

```
(BX900S1P00007-CB3) #saveconf ibswb-conf
Saving configuration into /usr/ftproot/ibswb-conf
Done.
(BX900S1P00007-CB3) #scp /usr/ftproot/ibswb-conf root@192.168.1.100:./ibswb-conf
root@192.168.1.100's password:
ibswb-conf                                         100% 4505      4.4KB/s   00:00
```

How to restore the configuration

1. Configure IP address, see the detail operation in [Configuration](#) below.
2. Copy *ibswb-conf* file to */usr/ftproot* of IB connection blade by *ftp* or *scp* command.
3. Execute *restoreconf* command from IB connection blade.

```
(BX900S1P00131-CB5) #scp root@192.168.1.100:./ibswb-conf /usr/ftproot
root@192.168.1.100's password:
ibswb-conf
(BX900S1P00131-CB5) #restoreconf ibswb-conf
Configuration was restored successfully
```



The configuration will be valid after IB connection blade reboots.

4.2 Configuration

In order to configure the function of IB connection blade using CLI (Command Line Interface), the user needs to either connect the management console to Management Blade and run a terminal software or connect to Management Blade via telnet.

After finishing configuring the network configuration (e.g. IP address), it is possible to login via telnet.

4.2.1 Configuration of LAN interface

1. Confirm that the prompt is displayed.

In case the prompt isn't displayed, press the **[Enter]** key to display the prompt.

Example : configure CB3

Press **[Ctrl]+[Q]** Return Console Menu
<BX900S1-CB3>config_net

2. Enter "config_net" and press the **[Enter]** key to display the menu below.

```
<BX900S1-CB3>config_net
Welcome to The Network Configuration Utility.
```

```
1) Configure Host Networking.
2) Configure IP Address.
3) Configuration for Host Networking and IP Address.
4) Exit from Network Configuration Utility.
Enter Your Selection:
```

3. Select "3)Configuration for Host Networking and IP Address." on the menu.

Enter Your Selection: 3

4. After displaying the following message, enter **[Y]**.

Do you want to change this configuration? [y/N]:y

5. After displaying the following message, enter **y** to enable "NETWORKING support".

Enable NETWORKING support [Y/n]:y

6. After displaying each following message, configure "Gateway Address" , "hostname" respectively.

If Gateway Address is "192.168.1.1" and hostname "ibsw1", enter the input values as follows.

Enter a Gateway Address:192.168.1.1

Enter a hostname [Ex: localhost.localdomain]:ibsw1

7. Check the configured values, and enter **y** to save the configuration.

Selected configuration:

NETWORKING=yes

GATEWAY=192.168.1.1

HOSTNAME=ibsw1

Do you want to save the selected configuration? [Y/n]:y

8. Enter "y" to display the IP address configuration menu.

Do you want to configure IP Address? [Y/n]:y

1) Configure by DHCP.

2) Static Configuration.

3) Exit from Configuration System Utility.

Enter Your Selection:

9. Select "2)Static Configuration." on the menu.

Enter Your Selection: 2

10. After displaying the following message, enter **y**.

Do you want to change this configuration? [y/N]:y

11. Configure "IP Address", "Netmask" respectively, then enter **y** when the message "Start Device On Boot?" is displayed.

If IP Address is "192.168.1.30" and Netmask "255.255.255.0", enter the input values as follows.

=====

Enter an IP Address:192.168.1.30

=====

Enter the Netmask:255.255.255.0

Start Device On Boot? [Y/n]:y

12. Check the configured values, and enter **y** to save the configuration.

DEVICE=eth0

IPADDR=192.168.1.30

NETMASK=255.255.255.0

HWADDR=00:02:C9:11:23:56

ONBOOT=yes

Do you want to save the selected configuration? [Y/n]:y

13. After displaying the following message, select "4)Exit from Network Configuration Utility." to finish the configuration.

```
=====
| Welcome to The Network Configuration Utility. |
=====
1) Configure Host Networking.
2) Configure IP Address.
3) Configuration for Host Networking and IP Address.
4) Exit from Network Configuration Utility.
Enter Your Selection: 4
```

14. Enter to enable the contents of configuration.

Do you want to enable the new network configuration? [Y/n]y

15. After displaying the prompt, enter "ifconfig" to check if the configured IP address is displayed.

```
<BX900S1-CB3>ifconfig
eth0      Link encap:Ethernet  HWaddr 00:02:C9:11:23:56
inet  addr:192.168.1.30  Bcast:192.168.14.255  Mask:255.255.255.0
          UP BROADCAST RUNNING MULTICAST  MTU:1500  Metric:1
          RX packets:9805  errors:0  dropped:0  overruns:0  frame:0
          TX packets:12  errors:0  dropped:0  overruns:0  carrier:0
          collisions:0  txqueuelen:1000
          RX bytes:1121747 (1.0 MiB)  TX bytes:601 (601.0 b)
Interrupt:25
```

4.2.2 Ethernet Connection via Management Blade

IB connection blade provides the function of Ethernet connection via the Management LAN port on a Management Blade.

Ethernet connection enables the user to connect to IB connection blade via telnet/ssh with CLI.

An example that an IP address 192.168.1.100/24 is assigned to management console and 192.168.1.10/24 to IB connection blade is illustrated below.

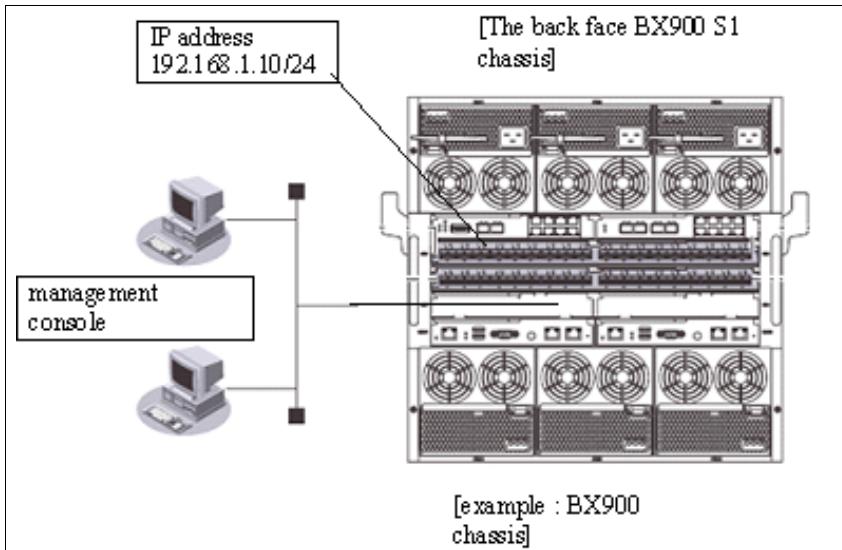


Figure 33: Ethernet connection

4.3 Basic Operation

This chapter describes the basic operation by the commands and how to get the information for the maintenance.

4.3.1 How to check System Information

This section describes how to check system information for this product.

The firmware version can be checked for this product.

swinfo command

This command shows

- GUID, InfiniScaleIV Firmware version,
- CPU Firmware version of the IB connection blade

```
<BX900S1-CB3>swinfo
<BX900S1-CB3 Information:
GUID: 0x0002c9020040c7a8
MAC: 00:02:c9:11:20:fc
IS4 Firmware revision: 7.2.326
CPU Firmware revision: Mellanox release mlnx405ex-1.0.2 build 2009-07-19
```

ifconfig command

This command shows and configures IP address, subnetmask, broadcast address.

The setting values by this command will disappear after IB connection blade reboots. If you would like to keep the values, you can use *config_net* command.

```
<BX900S1-CB3>ifconfig eth0 192.168.12.33/24
<BX900S1-CB3>ifconfig
eth0Link encap:Ethernet HWaddr 00:02:C9:11:20:FC
inet addr:192.168.12.33 Bcast:192.168.12.255 Mask:255.255.255.0
  UP BROADCAST RUNNING MULTICAST MTU:1500 Metric:1
  RX packets:5553 errors:0 dropped:0 overruns:0 frame:0
  TX packets:228 errors:0 dropped:0 overruns:0 carrier:0
  collisions:0 txqueuelen:1000
  RX bytes:745872 (728.3 KiB) TX bytes:39013 (38.0 KiB)
  Interrupt:25
loLink encap:Local Loopback
inet addr:127.0.0.1 Mask:255.0.0.0
  UP LOOPBACK RUNNING MTU:16436 Metric:1
  RX packets:28 errors:0 dropped:0 overruns:0 frame:0
```

4.3.2 SNMP Trap Settings

This chapter describes the SNMP trap settings for IB connection blade.

1. IP address setting

The IP address of IB connection blade is set up via Console Redirection of Management Blade.

2. Enabling SNMP Trap

SNMP Trap is disabled by default. How to enable the setting is as follows.

a) Command : *ibswsnmptrap*

b) Viewing the SNMP trap information (default)

```
(BX900S1P00007) #ibswsnmptrap
Trap is Disabled
Authentication Failure is Disabled
Address: 0.0.0.0
Community:
```

c) How to set

The details of a parameter of the *ibswsnmptrap* command are as follows:

```
(BX900S1P00007) #ibswsnmptrap --help
Usage: ibswsnmptrap [-t | enabled/disabled] [-a | 
enabled/disabled] [-i | IP address] [-c | Community]
-t : Send SNMP traps is enabled or disabled
-a : When SNMP trap is enabled, authentication
failure trap is enabled or disabled
-i : The destination of SNMP Trap receiver
-c : Set community for snmptraps
```

SNMP Trap send address: 192.168.11.253

Community name: public

Examples of the above settings are as follows:

```
(BX900S1P00007) #ibswsnmptrap -t enabled -i
192.168.11.253 -c public
(BX900S1P00007) #ibswsnmptrap
Trap is Enabled
Authentication Failure is Disabled
Address: 192.168.11.253
Community: public
```

3. Disabling SNMP Trap

How to disable the SNMP Trap is as follows:

```
(BX900S1P00007) #ibswsnmptrap -t disabled
```

```
(BX900S1P00007) #ibswsnmptrap
```

Trap is Disabled

Authentication Failure is Enabled

Address: 0.0.0.0

Community:

4.4 Users and Privileges

The CLI is entered via the Telnet/ssh interface of the management blade.

Three user levels are supported to execute CLI commands:

No	User name	Password	Authority level	Description
1	root	55ib-qdr-sw	root	Only Factory people and R&D use. The password is not usually used.
2	maintenance	#yrotcaf	maintenance	Our maintenance and system engineer use
3	ibsw	ibsw	user	For Read only users

Table 4: Default user



User maintenance runs as user *root* or *maintenance*.



Root password can be changed.



User names and passwords are enforced when logging-in via UART.



When CPU module boots up, Auto Stop is done by pressing Ctrl+B. Pressing Ctrl+B enters the boot console.

5 Troubleshooting

This chapter describes what to do if IB Mezzanine card and IB connection blade is not working properly.

If you cannot solve the problem, please contact the repair center:

<http://support.ts.fujitsu.com/com/support/index.html>

If you purchased the products from Fujitsu Japan, please consult your system engineer.

5.1 Check the Log File of OpenSM (Linux only)

Check the log file in the server (*/var/log/opensm.log*) that started up the Subnet Manager(OpenSM) as Master.

- ▶ Check list : Are any of the following logs contained in the */var/log/opensm.log*
 - SM port is down
 - Removed port with GUID:GUID number LID range [LID number, LID number] of node:
 - Reporting Generic Notice type:1 num:128 (Link state change) from LID:LID number GUID:GUID number

SM port is down

SubnetManager's master port was disconnected.

How to solve the problem: The following changes are required.

Please contact the repair center.

1. IB Mezzanine card

You need to change the IB Mezzanine card that ran as the Master of the SubnetManager.

For the method of how to specify the exact IB Mezzanine card, please refer to [section "The result of the command "ibnetdiscover"" on page 32](#).

2. ServerBlade

If you have changed IB Mezzanine card and you still see the same messages, you need to change the server blade that installed the IB Mezzanine card.

3. IB connection blade

If you have changed the above server blade and still get the same messages, you need to change IB connection blade. For details of how to specify the exact IB connection blade, please refer to section "[The result of the command "ibnetdiscover"](#)" on page 32.

Removed port with GUID:GUID number LID range [LID number, LID number] of node:

IB Mezzanine card, IB connection blade or IB cable was disconnected.

How to solve the problem: The following changes are required. Please contact the repair center.

1. IB Mezzanine card

You need to change the relevant IB Mezzanine card. You obtain the relevant IB Mezzanine card from GUID (or LID) information in the error messages and the output of *ibnetdiscover*.

For example:

```
Jul 30 23:34:18 600287 [4CD0D940] 0x02 ->
__osm_drop_mgr_remove_port: Removed port with
GUID:0x0002c90300041306 LID range [2, 2] of node:localhost
HCA-2
```

In this case, Error Card has the GUID 0x0002c90300041306 and the LID is 2 of Mezzanine Card).

2. IB connection blade

You need to change the relevant IB connection blade. You will get the relevant IB connection blade from GUID (or LID) information of the error messages and output of *ibnetdiscover*.

For example:

```
Jul 30 23:34:18 600424 [4CD0D940] 0x02 ->
__osm_drop_mgr_remove_port: Removed port with
GUID:0x0002c9020040c790 LID range [21, 21] of node:(none)
SW-1
```

In this case, Error IB connection blade has the GUID 0x0002c9020040c790 and the LID is 21.

For details of how to specify the relevant IB connection blade, refer to [section “The result of the command “ibnetdiscover”” on page 32](#).

3. IB cable

If you have changed the above IB connection blade and you still receive the messages, all IB cables connected to the IB connection blade need to be changed.

Reporting Generic Notice type:1 num:128 (Link state change) from LID:LID number GID:GID number: the port of IB connection blade was disconnected in the subnet.

How to solve the problem: The following changes are required. Please contact the repair center.

1. IB cable

You need to change the IB cable. You specify the IB cables by the following procedure.

- You specify the relevant IB connection blade from the LID information in the error messages.

For example:

```
Jul 31 22:56:46 264171 [42E93940] 0x02 ->
osm_report_notice: Reporting Generic Notice type:1
num:128 (Link state change) from LID:32
GID:fe80::2:c902:40:c790
```

In this case, LID 32's IB connection blade is corresponding.

For details of how to specify the relevant IB connection blade, refer to [section “The result of the command “ibnetdiscover”” on page 32](#)

- The target port is the one that removed the port number from original *ibnetdiscover* command output when the system was set up normally..

2. IB connection blade

If you have changed the above IB cable and still receive the same messages, you need to change the IB connection blade.

5.2 IB Mezzanine Card

You can confirm place and time the error occurred by referring to the following items on a server blade.

No.	Confirmation item
1	Active link LED for IB Mezzanine card on the front side of the blade
2	IB Mezzanine card driver message

Active link LED for IB Mezzanine card

When an active link LED for IB Mezzanine card on the front side of the server blade goes dark, you can confirm the place where the error occurred using the following confirmation contents.

No.	Confirmation item	Judgment	To do
1	Is the server blade turned on?	YES	Go to No.2
		NO	Turn on the server blade.
2	Are the LEDs of all server blades turned off?	YES	Confirm OpenSM running on the MASTER OpenSM server.
		NO	Go to No.3
3	Is the LED of one server blade in the system turned off?	YES	Go to No.4
		NO	Go to No.5
4	Is the IB Mezzanine card mounted rigidly?	YES	<ol style="list-style-type: none"> 1. Replace the IB Mezzanine card. 2. If the LED is still turned off after replacing the IB Mezzanine card, replace the server blade. 3. If the LED is still turned off after replacing the server blade, replace the IB connection blade. 4. If the LED is still turned off, please consult technical.support@ts.fujitsu.com <p>If you purchased the products from Fujitsu Japan, please consult the system engineer.</p>
		NO	Mount the IB Mezzanine card more rigidly.
		NO	
5	Does IB connection blade work normally? Please refer to section “ IB Connection Blade ” on page 76.	YES	Go to No.4
		NO	<p>Replace the IB connection blade. Please consult technical.support@ts.fujitsu.com</p> <p>If you purchased the products from Fujitsu Japan, please consult the system engineer.</p>

Table 5: Indication lamps

IB Mezzanine card driver messages (Linux)

Confirm the error message of IB Mezzanine card driver by referring to the `/var/log/messages` log file. You refer to the log file using the text editor etc.

No.	Classification	To do
1	Installed FW has unsupported command interface revision %d. (Installed FW version is %d.%d.%03d) This driver version supports only revisions %d to %d.	The IB Mezzanine card must be replaced. Please consult technical.support@ts.fujitsu.com
2	This driver version supports only revisions %d to %d.	If you purchased the products from Fujitsu Japan, please consult the system engineer.
3	Internal error detected: buf[%02x]: %08x	
4	Failed to initialize user access region table, aborting.	
5	Couldn't map EQ doorbell for EQN 0x%06x	
6	Couldn't map interrupt clear register, aborting.	
7	Couldn't allocate FW area, aborting.	(note1)
8	Failed to map MCG context memory, aborting.	

Table 6: Error messages of IB Mezzanine card driver (Red Hat Enterprise Linux)

No.	Classification	To do
9	Failed to allocate driver access region, aborting.	
10	Failed to initialize memory region table, aborting.	
11	Failed to initialize event queue table, aborting.	
12	Failed to switch to event-driven firmware commands, aborting.	
13	NOP command failed to generate interrupt (IRQ %d), aborting.	The IB Mezzanine card must be replaced. Please consult technical.support@ts.fujitsu.com
14	Failed to initialize completion queue table, aborting.	If you purchased the products from Fujitsu Japan, please consult the system engineer.
15	Failed to initialize shared receive queue table, aborting.	
16	Failed to initialize queue pair table, aborting.	
17	Failed to initialize multicast group table, aborting.	
18	Failed to reset Mezzanine card, aborting.	
19	Failed to init command interface, aborting.	
20	Only %d UAR pages (need more than 128)	(note1)
21	Profile requires 0x%llx bytes; won't fit in 0x%llx bytes of context memory.	(note1)

Table 6: Error messages of IB Mezzanine card driver (Red Hat Enterprise Linux)

No.	Classification	To do
22	Couldn't allocate memory to save Mezzanine card PCI header, aborting.	
23	Couldn't save Mezzanine card PCI header, aborting.	
24	Couldn't map Mezzanine card reset register, aborting.	
25	PCI device did not come back after reset, aborting.	The IB Mezzanine card must be replaced. Please consult technical.support@ts.fujitsu.com
26	Couldn't restore Mezzanine card PCI Express Device Control register, aborting.	
27	Couldn't restore Mezzanine card PCI Express Link control register, aborting.	
28	Couldn't restore Mezzanine card reg %x, aborting.	
29	Couldn't restore Mezzanine card COMMAND, aborting.	If you purchased the products from Fujitsu Japan, please consult the system engineer.

Table 6: Error messages of IB Mezzanine card driver (Red Hat Enterprise Linux)

note1: When a message occurs frequently, confirm the following:

No.	Confirmation contents	Judgment	To do
1	Is the IB Mezzanine card mounted rigidly?	YES	The IB Mezzanine card must be replaced. When the message is displayed after the IB Mezzanine cards are exchanged, it is necessary to replace the server blade.
		NO	Place the IB Mezzanine card again.

Work after IB Mezzanine card has been replaced (Linux)

1. The server blade is mounted on the main body of BX900 and the power supply is turned on.
2. Linux is started, */sbin/lspci* command is executed by the root account, and it is confirmed that the IB Mezzanine card is recognized.

Execution result:

```
10:00.0 InfiniBand: Mellanox Technologies MT26428 [ConnectX IB QDR, PCIe 2.0
5GT/s] (rev a0) Subsystem: Mellanox Technologies MT26428 [ConnectX IB
QDR, PCIe 2.0 5GT/s]
30:00.0 InfiniBand: Mellanox Technologies MT26428 [ConnectX IB QDR, PCIe 2.0
5GT/s] (rev a0)
Subsystem: Mellanox Technologies MT26428 [ConnectX IB QDR, PCIe 2.0 5GT/s]~
```

The enhancing board slot where this card is mounted can be confirmed by the value of Bus#.

10:00.0 -- Enhancing board slot 1

30:00.0 -- Enhancing board slot 2

3. Refer to [section “Network configuration diagram” on page 29](#), update a GUID table.

IB Mezzanine card driver messages (Windows Server 2008 R2)

Confirm the message of the IB Mezzanine card driver by referring to the event log. You refer to the event log using the Event Viewer.

Event ID Error Level	Classification	To do
0x0041 Error	xxx: Failed to open Channel Adapter.	
0x0042 Error	xxx: Failed to allocate Protection Domain.	
0x0043 Error	xxx: Failed to create receive Completion Queue.	
0x0044 Error	xxx: Failed to create send Completion Queue.	
0x0045 Error	xxx: Failed to create Queue Pair.	
0x0046 Error	xxx: Failed to get Queue Pair number.	The IB Mezzanine card must be replaced.
0x0047 Error	xxx: Failed to create DMA Memory Region.	Please consult technical.support@ts.fujitsu.com
0x0048 Error	xxx: Failed to create receive descriptor pool.	If you purchased the products from Fujitsu Japan, please consult the system engineer.
0x0049 Error	xxx: Failed to create NDIS_PACKET pool to receive indications.	
0x004A Error	xxx: Failed to create NDIS_BUFFER pool to receive indications.	
0x004B Error	xxx: Failed to create NDIS_PACKET pool to send processing.	
0x004C Error	xxx: Failed to create NDIS_BUFFER pool to send processing.	

Table 7: [source:ipoib]

Event ID Error Level	Classification	To do
0x004D Error	xxx: Failed to allocate receive indication array.	
0x004E Error	xxx: Subnet Administrator query for port information timed out. Make sure the SA is functioning properly. Increasing the number of retries and retry timeout adapter parameters may solve the problem.	
0x004F Error	xxx: Subnet Administrator failed the query for port information. Make sure the SA is functioning properly and compatible.	The IB Mezzanine card must be replaced.
0x0050 Error	xxx: Subnet Administrator query for port information failed.	Please consult technical.support@ts.fujitsu.com
0x0055 Error	xxx: Subnet Administrator failed query for broadcast group information.	If you purchased the products from Fujitsu Japan, please consult the system engineer.
0x0056 Error	xxx: Subnet Administrator failed request to joining broadcast group.	
0x0057 Error	xxx: The local port rate is too slow for the existing broadcast MC group.	
0x0058 Error	xxx: Incorrect value or non-existing registry for the required IPoIB parameter %3, overriding it by default value: %4	

Table 7: [source:ipoib]

Event ID Error Level	Classification	To do
0x005B Error	xxx: Pkey index not found for partition, change switch pkey configuration.	
0x005C Error	xxx: Connected Mode failed to initialize, disabled. Interface will use default UD QP transport.	The IB Mezzanine card must be replaced.
0x005E Error	xxx: SetDeviceRegistrationAttributes failed.	Please consult technical.support@ts.fujitsu.com
0x005F Error	xxx: SetAdapterRegistrationAttributes failed.	
0x0060 Error	xxx: SetOffloadAttributes failed.	
0x0061 Error	xxx: ipoib_create_adapter failed.	
0x0062 Error	xxx : ipoib_start_adapter failed.	

Table 7: [source:ipoib]

Event ID Error Level	Classification	To do
0x0003 Error	The message changes in each event.	
0x0004 Error	mlx4_bus has started in non-operational mode.	
0x0005 Error	<p>mlx4_bus has failed to start even in non-operational mode.%n</p> <p>Look into the the previous error messages.</p>	
0x0007 Error	<p>MAP_FA command failed with error %2.%n</p> <p>The adapter card is non-functional.%n</p> <p>Most likely a FW problem.%n</p> <p>Please burn the last FW and restart the mlx4_bus driver.</p>	<p>The IB Mezzanine card must be replaced.</p> <p>Please consult</p> <p>technical.support@ts.fujitsu.com</p>
0x0008 Error	<p>RUN_FW command failed with error %2.%n</p> <p>The adapter card is non-functional.%n</p> <p>Most likely a FW problem.%n</p> <p>Please burn the last FW and restart the mlx4_bus driver.</p>	<p>If you purchased the products from Fujitsu Japan, please consult the system engineer.</p>
0x0009 Error	<p>QUERY_FW command failed with error %2.%n</p> <p>The adapter card is non-functional.%n</p> <p>Most likely a FW problem.%n</p> <p>Please burn the last FW and restart the mlx4_bus driver.</p>	

Table 8: [source: mlx4_bus]

Event ID Error Level	Classification	To do
0x000B Error	<p>QUERY_DEV_CAP command failed with error %2.%n</p> <p>The adapter card is non-functional.%n</p> <p>Most likely a FW problem.%n</p> <p>Please burn the last FW and restart the mlx4_bus driver.</p>	
0x000C Error	<p>QUERY_ADAPTER command failed with error %2.%n</p> <p>The adapter card is non-functional.%n</p> <p>Most likely a FW problem.%n</p> <p>Please burn the last FW and restart the mlx4_bus driver.</p>	<p>The IB Mezzanine card must be replaced.</p> <p>Please consult technical.support@ts.fujitsu.com</p>
0x000D Error	<p>Too few QPs were requested (requested %2, reserved for FW %3).%n</p> <p>The adapter card is non-functional.%n</p> <p>Please increase the Registry LogNumQp parameter under HKLM\System\CurrentControlSet\Services\mlx4_bus\Parameters.</p>	<p>If you purchased the products from Fujitsu Japan, please consult the system engineer.</p>

Table 8: [source: mlx4_bus]

Event ID Error Level	Classifi- cation	To do
0x0003 Error	The message changes in each event.	<p>The IB Mezzanine card must be replaced.</p> <p>Please consult technical.support@ts.fujitsu.com</p> <p>If you purchased the products from Fujitsu Japan, please consult the system engineer.</p>

Table 9: [source:ibbus]

Work after IB Mezzanine card has been exchanged (Windows Server 2008 R2)

1. Turn on the server.
2. Make sure that the InfiniBand devices are recognized properly in Device Manager.

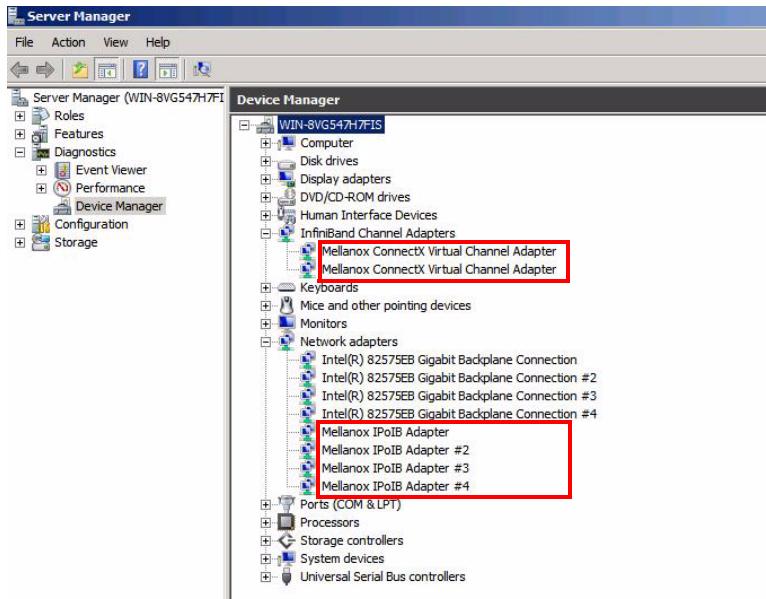


Figure 34: Device manager

3. Refer to [section “Configuration sheet” on page 29](#), update a GUID table.

5.3 IB Connection Blade

IB Connection Blade Message

Connect to IB connection blade via MMB (Console Redirection). The reference to the */var/log/messages* log file uses the text editor etc.

No.	Classification	To do
1	Installed FW has unsupported command interface revision %d. (Installed FW version is %d.%d.%03d) This driver version supports only revisions %d to %d.	The IB connection blade must be replaced. Please consult technical.support@ts.fujitsu.com
2	This driver version supports only revisions %d to %d.	If you purchased the products from Fujitsu Japan, please consult the system engineer.
3	Internal error detected: buf[%02x]: %08x	

Table 10: IB connection blade messages

snmptrap messages

When SNMP Trap of the IB connection blade is set to *enable*, following snmp trap is transmitted to SNMP Manager.

No.	SNMP Trap	OID	Content and Action
1	ASIC chip down	1.3.6.1.4.1. 33049.2.2.1.2.1	[Content] The ASIC of the IB connection blade was downed. [Action] Reboot of the IB connection blade. When it doesn't restore, even if rebooted, the IB connection blade must be replaced. When trap is frequent even if restore after Reboot, the IB connection blade must be replaced. Please consult technical.support@ts.fujitsu.com If you purchased the products from Fujitsu Japan, please consult the system engineer.

Table 11: snmp trap

No.	SNMP Trap	OID	Content and Action
2	ASIC overtemp	1.3.6.1.4.1. 33049.2.2.1.2.2	[Content] The ASIC of the IB connection blade detected temperature abnormality. [Action] Confirm that ambient temperature of BX900 or BX400 system does not have abnormality. When ambient temperature does not have abnormality, the IB connection blade must be replaced. Please consult technical.support@ts.fujitsu.com If you purchased the products from Fujitsu Japan, please consult the system engineer.
3	Internal I2C error	1.3.6.1.4.1. 33049.2.1.2.1	[Content] The Internal I2C error occurred. [Action] Reboot of the IB connection blade. If it doesn't restore even if rebooted, the IB connection blade must be replaced. When trap is frequent even if restore after reboot, the IB connection blade must be replaced. Please consult technical.support@ts.fujitsu.com If you purchased the products from Fujitsu Japan, please consult the system engineer.
4	cold start	MIB-II standard trap	[Content] The cold start occurred. [Action] When cold start not intended happens frequently, the IB connection blade must be replaced. Please consult technical.support@ts.fujitsu.com If you purchased the products from Fujitsu Japan, please consult the system engineer.

Table 11: snmp trap

Work after the exchange

1. IB connection blade is mounted on the main body of BX900/BX400 and the power supply is turned on.
2. Confirm that the Status LED of the IB connection blade lights to green. It takes about 90 seconds till the CPU of the IB connection blade starts.
3. Connect to the console of the IB connection blade via telnet and wait for the prompt of the IB connection blade to be displayed. When no prompt is displayed, push *Enter* and confirm that a prompt is displayed.
4. Set ip address which there was it before exchange.

Example:

```
#ifconfig eth0 192.168.1.100/24
```

5. Restore composition definition information. Please refer to [section “Save and restore the configuration” on page 52](#).
6. Connect IB-Cable to the same port as before.
7. Confirm that LED of P and L/A lights. It might take about one minute to lighting LED.
8. *ibstat* is executed from the console of the IB connection blade, and comfirm that State: Active, Physical state: LinkUp, Rate: 40, Base lid: figures other than 0 is displayed.

```
<BX900S1-CB5>ibstat
Switch 'is4_0'
Switch type: MT48436
Number of ports: 0
Firmware version: 7.2.326
Hardware version: a0
Node GUID: 0x0002c9020040bbbb
System image GUID: 0x0002c9020040bbbb
Port 0:
State: Active
Physical state: LinkUp
Rate: 40
Base lid: 28
LMC: 0
SM lid: 16
Capability mask: 0x02500848
Port GUID: 0x0002c9020040bbbb
```

9. Refer to [section “Network Configuration Information” on page 28](#), update a GUID table.

Index

A

Auto Stop [60](#)

B

BX400 S1 [5](#)

BX900 S1 [5](#)

C

CLI [60](#)

Command line interface [60](#)

configuration files [19](#)

configuration sheet [29](#)

CPU module

boot [60](#)

D

disk Space

for installation [16](#)

disk space

for Windows installation [37](#)

E

ethernet connection [57](#)

G

GUID [35](#)

I

ibnetdiscover [29](#)

ibstat [28](#)

ifconfig [58](#)

InfiniBand connection blade
connection [51](#)

installer privileges [16](#)

Windows [37](#)

K

kernel [18](#)

kernel modules [19](#)

L

LEDs [12](#)

Linux [16](#)

log file [61](#)

M

Management Blade [51](#)

login [51](#)

master [28](#)

memory limitations [19](#)

Mezzanine card

exchange [68](#)

GUID [35](#)

hardware installation [5](#)

technical data [13](#)

N

network configuration diagram [28](#)

notational conventions [5](#)

O

OFED [5](#)

SVIM [20](#)

OpenFabrics Enterprise

Distribution [5](#)

OpenSM [28, 61](#)

installation [18](#)

opensm [20](#)

operating system [16](#)

P

PRIMERGY BX400 [5](#)

PRIMERGY BX900 [5](#)

PRIMERGY BX900 Blade Server [8](#)

Q

QSFP [11](#)

Quad Small Form factor

Pluggable [11](#)

R

Red Hat Enterprise Linux [17](#)

repair center [61](#)

Root password [60](#)

Index

S

Server View Installation Manager [20](#)

SM [28](#)

software requirements

 Windows [37](#)

standby [29](#)

Subnet Manager [28](#)

subnet manager [20, 27](#)

 OpenSM [61](#)

support [61](#)

SVIM [20](#)

swinfo [58](#)

switch LEDs [12](#)

T

technical support [61](#)

telnet connection [52](#)

U

user maintenance [60](#)

W

Windows installation [37](#)